

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/642,452	08/18/2000	Josef Bauer	POO,1701	7124

7590 10/28/2004

Morrison & Foerster LLP  
1650 Tysons Boulevard  
Suite 300  
McLean, VA 22102

EXAMINER

LERNER, MARTIN

ART UNIT	PAPER NUMBER
----------	--------------

2654

DATE MAILED: 10/28/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

**Office Action Summary**

Application No.

09/642,452

Applicant(s)

BAUER ET AL.

Examiner

Martin Lerner

Art Unit

2654

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
  - If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
  - If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
  - Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1) ☒ Responsive to communication(s) filed on 16 September 2004.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4) ☒ Claim(s) 15 to 28 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 15 to 26 and 28 is/are rejected.
- 7) ☒ Claim(s) 27 is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 16 July 2004 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
- Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☒ All b) ☐ Some \* c) ☐ None of:
1. ☒ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)  
Paper No(s)/Mail Date \_\_\_\_\_
- 4) ☐ Interview Summary (PTO-413)  
Paper No(s)/Mail Date. \_\_\_\_\_
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other: \_\_\_\_\_

## DETAILED ACTION

### ***Claim Rejections - 35 USC § 102***

1. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

2. Claims 15, 16, 20 to 23, and 28 are rejected under 35 U.S.C. 102(b) as being anticipated by *Power et al.*

Regarding independent claims 15 and 28, *Power et al.* discloses a speech recognition method and system, comprising:

“performing a preliminary speech recognition of a voice signal to segment the voice signal into words and pauses and converting the words into text” – classifier 34 receives successive feature vectors and operates on each with a plurality of models corresponding to different words, phonemes or phrases to generate recognition results (column 4, lines 54 to 57: Figure 2); classifier 34 comprises a classifying process 341 and a state memory 342; a state field is provided for noise/silence state at the beginning of a word and state field for a noise/silence state at the end of a word (column 5, lines 24 to 37: Figure 3); implicitly, recognition of particular words implies “converting the words into text”;

“determining an average silence volume during the pauses” – pause detector 37 comprises a SNR detector 372 to read an average energy level buffer 376 to determine

Art Unit: 2654

a representative energy level over the frames currently identified as speech and a representative energy level over frames currently represented as noise; the representative measure comprises the mean running energy level running over the noise segments (column 8, lines 20 to 32: Figure 10);

“determining an average word volume for the words” – pause detector 37 comprises a SNR detector 372 to read an average energy level buffer 376 to determine a representative energy level over the frames currently identified as speech and a representative energy level over frames currently represented as noise; the representative measure comprises a peak average energy level over the speech segment (column 8, lines 20 to 32: Figure 10);

“calculating a difference between the average word volume and the average silence volume” – a signal to noise ratio value, SNR, is calculated (column 8, lines 33 to 33 to 39: Figures 10 and 12); a signal to noise ratio, SNR, represents a “difference” because a ratio compares the relative magnitudes of two quantities;

“evaluating a word, having a volume difference between the average word volume and the average silence volume is lower than a predetermined threshold, as having been incorrectly recognized” – rejector 36 is arranged to test the confidence of identification of a word by parser 35; if the identification is suspect, it is rejected; silence is detected by testing whether the SNR calculated by the SNR detector 373 lies below a very low threshold; the tests performed by the rejector include a test using the signal to noise ratio calculated by SNR detector 372 to reject noisy conditions and out-of-vocabulary words (column 9, line 51 to column 10, line 43: Figures 2 and 17).

Regarding claim 16, it is implicit that signal to noise ratios are measured in decibels; a decibel is a logarithm of a signal energy.

Regarding claim 20, *Power et al.* discloses thresholds, but does not say that the thresholds are adjusted or adapted; thus, the thresholds are constants.

Regarding claim 21, *Power et al.* discloses a recognition rejector 36 is arranged to reject recognition of a word recognized by parser 35 if recognition is unreliable (column 4, lines 61 to 66: Figure 2).

Regarding claims 22 and 23, *Power et al.* discloses rejector 36 issues a "query" signal which enables the utilizing apparatus 4 to initiate a confirmatory dialogue by synthesizing a phrase asking the user to repeat the word (column 9, lines 52 to 61); a confirmatory dialogue is "a message".

### ***Claim Rejections - 35 USC § 103***

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

4. Claim 24 is rejected under 35 U.S.C. 103(a) as being unpatentable over *Power et al.* in view of *Polikaitis et al.*

*Power et al.* discloses a recognition rejector 36 arranged to reject recognition of a word recognized by parser 35 if recognition is unreliable, wherein a rejection signal is output from the rejector 36 to control signal output 38 for use in controlling a utilizing

Art Unit: 2654

apparatus 4 (column 4, line 61 to column 5, line 2: Figures 1 and 2). However, *Power et al.* does not specifically provide a message output for a user to speak louder so that an adequate distance is achieved between the average word volume and the average silence volume. *Polikaitis et al.* teaches a speech recognition device for screening speech input, wherein error procedures are performed if ratios of speech energy are less than various thresholds. Specifically, *Polikaitis et al.* discloses if Control4 is option A, the user is prompted in step 270 to repeat the voice instruction and is prompted to speak louder (column 9, lines 5 to 8: Figure 2). Implicitly, speaking louder causes SpeechEnergy ("average word volume") to increase relative to NoiseEnergy ("average silence volume") as an increased signal-to-noise ratio ("so that an adequate distance is achieved"). *Polikaitis et al.* teaches an objective of screening speech input so that a speech recognizer operates correctly because speech recognition technology does not work well when the user speaks too softly. (Column 1, Line 39 to Column 2, Line 12) It would have been obvious to one having ordinary skill in the art to provide a message to a user to speak louder as taught by *Polikaitis et al.* in the speech recognition system with sequence parsing and rejection of *Power et al.* for the purpose of screening speech input so that a recognizer operates correctly when the user speaks too softly.

5. Claims 17 to 19 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Power et al.* in view of *Wu et al.*

Regarding claim 17, *Power et al.* does not disclose adapting thresholds on the basis of the global difference, although adaptive thresholds are fairly well known. *Wu et*

Art Unit: 2654

*al.* teaches a generally similar speech recognition method for analyzing endpoints in speech with signal-to-noise ratios, where speech recognition is only performed if a predetermined restart threshold level is identified. (Column 9, Line 56 to Column 10, Line 5) *Wu et al.* employs adaptive thresholds,  $T_s$ ,  $T_e$ ,  $T_{sr}$ ,  $T_{er}$ , defined in terms of an average background noise level  $N_{bg}$ , and average speech energy levels,  $E_{ls}$  and  $E_{le}$ . (Column 7, Line 25 to Column 9, Line 31: Figures 8, 9(a) and 9(b)) Specifically, *Wu et al.* says the method is advantageous for eliminating errors due to mistaking breathing for actual speech. (Column 9, Line 56 to Column 10, Line 5) It would have been obvious to one having ordinary skill in the art to employ adaptive thresholds defined in term of average speech energy and average noise energy as suggested by *Wu et al.* for the thresholds of *Power et al.* in order to eliminate errors due to mistaking breathing for actual speech.

Regarding claim 18 *Wu et al.* discloses the thresholds are related to the signal-to-noise ratios, defined in terms of differences  $E_{ls} - N_{bg}$  and  $E_{le} - N_{bg}$  (column 8, lines 24 to 65).

Regarding claim 19, *Wu et al.* discloses general formulae for adaptive thresholds  $T_{sr}$  and  $T_{er}$ , where the thresholds are diminished by a factor  $-c_3 N_{bg}$ , and  $c_3$  is a constant to account for conditions of unstable background noise (column 9, lines 20 to 31).

6. Claims 25 and 26 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Power et al.* in view of *Polikaitis et al.* as applied to claim 24 above, and further in view of *Wu et al.*



Regarding claim 25, *Power et al.* discloses thresholds, but does not determine an average silence volume for an individual pause. However, *Wu et al.* teaches a generally similar speech recognition method for analyzing endpoints in speech with signal-to-noise ratios, where speech recognition is only performed if a predetermined restart threshold level is identified. (Column 9, Line 56 to Column 10, Line 5) *Wu et al.* determines an average background noise level  $N_{bg}$  on the basis of segments of silence energy defining a reliable island. (Column 7, Lines 25 to 42: Figure 8) Similarly, *Wu et al.* determines average speech energy levels,  $E_{ls}$  and  $E_{le}$ , on the basis of segments of speech energy defining a reliable island. (Column 7, Line 58 to Column 8, Line 23: Figures 9(a) and 9(b)). *Wu et al.* says the method is advantageous for eliminating errors due to mistaking breathing for actual speech. (Column 9, Line 56 to Column 10, Line 5) It would have been obvious to one having ordinary skill in the art to determine a difference between average speech energy and average noise energy in terms of immediately preceding or immediately following pauses as suggested by *Wu et al.* instead of the average noise energy and peak average speech energy of *Power et al.* for the purpose of eliminating errors due to mistaking breathing for actual speech.

Regarding claim 26, *Power et al.* discloses thresholds, but does not expressly average silence volume over a plurality of successive pauses to determine a difference between average word volume and average silence volume. However, *Wu et al.* determines an average background noise level,  $N_{bg}$ , on the basis of segments of silence energy defining a reliable island, and similarly, determines average speech energy levels,  $E_{ls}$  and  $E_{le}$ , on the basis of segments of speech energy defining a reliable island.

Art Unit: 2654

(Column 7, Line 25 to Column 8, Line 23: Figures 8, 9(a), and 9(b)). *Wu et al.* says the method is advantageous for eliminating errors due to mistaking breathing for actual speech. (Column 9, Line 56 to Column 10, Line 5) It would have been obvious to one having ordinary skill in the art to combine the segmental energy averaging method of *Wu et al.* for the average noise energy and peak average speech energy of *Power et al.* so as to determine the global average silence energy on the basis of a sum of the energies of successive silence segments for the purpose of eliminating errors due to mistaking breathing for actual speech.

#### ***Allowable Subject Matter***

7. Claim 27 is objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

#### ***Response to Arguments***

8. Applicants' arguments filed 16 September 2004 have been considered but are moot in view of the new grounds of rejection.

#### ***Conclusion***

9. The prior art made of record and not relied upon is considered pertinent to Applicants' disclosure.

Shaffer et al. discloses related art.

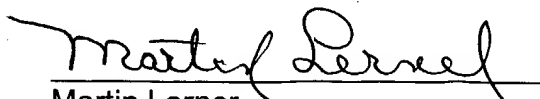
Art Unit: 2654

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Martin Lerner whose telephone number is (703) 308-9064. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (703) 305-9645. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

ML  
10/18/04

  
Martin Lerner  
Examiner  
Group Art Unit 2654